0280

RAW SEQUENCE LISTING   DATE: 10/19/2000
PATENT APPLICATION: US/09/676,718   TIME: 11:24:31

Input Set : A:\56113.app
Output Set: N:\CRF3\10192000\I676718.raw

ENTERED

```
    3 <110> APPLICANT: Gladyshev, V. et al.
    5 <120> TITLE OF INVENTION: Mammalian selenoprotein differentially expressed in
    6       tumor cells
    8 <130> FILE REFERENCE: 56113
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/676,718
C--> 11 <141> CURRENT FILING DATE: 2000-09-28
   13 <150> PRIOR APPLICATION NUMBER: 60/080,850
   14 <151> PRIOR FILING DATE: 1998-04-06
   16 <150> PRIOR APPLICATION NUMBER: PCT/US99/07560
   17 <151> PRIOR FILING DATE: 1999-04-06
   19 <160> NUMBER OF SEQ ID NOS: 15
   21 <170> SOFTWARE: PatentIn Ver. 2.0
   23 <210> SEQ ID NO: 1
   24 <211> LENGTH: 162
   25 <212> TYPE: PRT
   26 <213> ORGANISM: Homo sapiens
   28 <220> FEATURE:
   29 <221> NAME/KEY: SITE
   30 <222> LOCATION: (93)
   31 <223> OTHER INFORMATION: Xaa represents selenocysteine
   33 <400> SEQUENCE: 1
   34 Met Ala Ala Gly Pro Ser Gly Cys Leu Val Pro Ala Phe Gly Lys Arg
   35   1               5                  10                  15
   37 Leu Leu Leu Ala Thr Val Leu Gln Ala Val Ser Ala Phe Gly Ala Glu
   38              20                  25                  30
   40 Phe Ser Ser Glu Ala Cys Arg Glu Leu Gly Phe Ser Ser Asn Leu Leu
   41              35                  40                  45
   43 Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe Asn Leu Leu Gln Leu Asp
   44         50                  55                  60
   46 Pro Asp Cys Arg Gly Cys Cys Gln Glu Glu Ala Gln Phe Glu Thr Lys
   47   65                  70                  75                  80
   49 Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val Cys Gly Xaa Lys Leu Gly
   50              85                  90                  95
   52 Arg Phe Pro Gln Val Gln Ala Phe Val Arg Ser Asp Lys Pro Lys Leu
   53             100                 105                 110
   55 Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg Gly Ser Asp Pro Val Leu
   56             115                 120                 125
   58 Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala Glu Glu Leu Ser Ile Leu
   59         130                 135                 140
   61 Lys Trp Asn Thr Asp Ser Val Glu Glu Phe Leu Ser Glu Lys Leu Glu
   62 145                 150                 155                 160
   64 Arg Ile
   68 <210> SEQ ID NO: 2
   69 <211> LENGTH: 1244
   70 <212> TYPE: DNA
   71 <213> ORGANISM: Homo sapiens
   73 <220> FEATURE:
```

RAW SEQUENCE LISTING DATE: 10/19/2000
PATENT APPLICATION: US/09/676,718 TIME: 11:24:31

Input Set : A:\56113.app
Output Set: N:\CRF3\10192000\I676718.raw

```
 74 <221> NAME/KEY: CDS
 75 <222> LOCATION: (5)..(493)
 77 <220> FEATURE:
 78 <221> NAME/KEY: misc_feature
 79 <222> LOCATION: (281)..(283)
 80 <223> OTHER INFORMATION: TGA codon codes for selenocysteine, Xaa
 82 <400> SEQUENCE: 2
 83 agcg atg gcg gct ggg ccg agt ggg tgt ctg gtg ccg gcg ttt ggg cta   49
 84      Met Ala Ala Gly Pro Ser Gly Cys Leu Val Pro Ala Phe Gly Leu
 85       1            5                  10                  15
 87 cgg ttg ttg ttg gcg act gtg ctt caa gcg gtg tct gct ttt ggg gca   97
 88 Arg Leu Leu Leu Ala Thr Val Leu Gln Ala Val Ser Ala Phe Gly Ala
 89           20                  25                  30
 91 gag ttt tca tcg gag gca tgc aga gag tta ggc ttt tct agc aac ttg  145
 92 Glu Phe Ser Ser Glu Ala Cys Arg Glu Leu Gly Phe Ser Ser Asn Leu
 93           35                  40                  45
 95 ctt tgc agc tct tgt gat ctt ctc gga cag ttc aac ctg ctt cag ctg  193
 96 Leu Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe Asn Leu Leu Gln Leu
 97           50                  55                  60
 99 gat cct gat tgc aga gga tgc tgt cag gag gaa gca caa ttt gaa acc  241
100 Asp Pro Asp Cys Arg Gly Cys Cys Gln Glu Glu Ala Gln Phe Glu Thr
101       65                  70                  75
103 aaa aag ctg tat gca gga gct att ctt gaa gtt tgt gga tga aaa ttg  289
W--> 104 Lys Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val Cys Gly Xaa Lys Leu
105  80                  85                  90                  95
107 gga agg ttc cct caa gtc caa gct ttt gtt agt gat aaa ccc aaa  337
108 Gly Arg Phe Pro Gln Val Gln Ala Phe Val Arg Ser Asp Lys Pro Lys
109               100                 105                 110
111 ctg ttc aga gga ctg caa atc aag tat gtc cgt ggt tca gac cct gta  385
112 Leu Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg Gly Ser Asp Pro Val
113           115                 120                 125
115 tta aag ctt ttg gac gac aat ggg aac att gct gaa gaa ctg agc att  433
116 Leu Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala Glu Glu Leu Ser Ile
117           130                 135                 140
119 ctc aaa tgg aac aca gac agt gta gaa gaa ttc ctg agt gaa aag ttg  481
120 Leu Lys Trp Asn Thr Asp Ser Val Glu Glu Phe Leu Ser Glu Lys Leu
121       145                 150                 155
123 gaa cgc ata taa atcttgctta aattttgtcc tatccttttg ttaccttatc  533
124 Glu Arg Ile
125 160
127 aaatgaaata ttacagcacc tagaaaataa tttagttttg cttgcttcca ttgatcagtc 593
129 ttttacttga ggcattaaat atctaattaa atcgtgaaat ggcagtatag tccatgatat 653
131 ctaaggagtt ggcaagctta acaaaaccca ttttttataa atgtccatcc tcctgcattt 713
133 gttgatacca ctaacaaaat gctttgtaac agacttgcgg ttaattatgc aaatgatagt 773
135 ttgtgataat tggtccagtt ttacgaacaa cagatttcta aattagagag gttaacaaga 833
137 cagatgatta ctatgcctca tgtgctgtgt gctctttgaa aggaatgaca gcagactaca 893
139 aagcaaataa gatatactga gcctcaacag attgcctgct cctcagagtc tctcctattt 953
141 ttgtattacc cagctttctt tttaatacaa atgttattta tagtttacaa tgaatgcact 1013
143 gcataaaaac tttgtagctt cattattgta aaacatattc aagatcctac agtaagagtg 1073
```

```
145 aaacattcac aaagatttgc gttaatgaag actacacaga aaacctttct agggatttgt 1133
147 gtggatcaga tacatacttg gcaaattttt gagttttaca ttcttacaga aaagtccatt 1193
149 taaaagtgat catttgtaag accaaaaatat aaataaaaag tttcaaaaat c          1244
152 <210> SEQ ID NO: 3
153 <211> LENGTH: 489
154 <212> TYPE: DNA
155 <213> ORGANISM: Homo sapiens
157 <220> FEATURE:
158 <221> NAME/KEY: CDS
159 <222> LOCATION: (1)..(489)
161 <220> FEATURE:
162 <221> NAME/KEY: misc_feature
163 <222> LOCATION: (277)..(279)
164 <223> OTHER INFORMATION: TGA codon codes for Selenocysteine, Xaa
166 <400> SEQUENCE: 3
167 atg gcg gct ggg ccg agt ggg tgt ctg gtg ccg gcg ttt ggg cta cgg  48
168 Met Ala Ala Gly Pro Ser Gly Cys Leu Val Pro Ala Phe Gly Leu Arg
169  1               5                  10                  15
171 ttg ttg ttg gcg act gtg ctt caa gcg gtg tct gct ttt ggg gca gag  96
172 Leu Leu Leu Ala Thr Val Leu Gln Ala Val Ser Ala Phe Gly Ala Glu
173             20                  25                  30
175 ttt tca tcg gag gca tgc aga gag tta ggc ttt tct agc aac ttg ctt 144
176 Phe Ser Ser Glu Ala Cys Arg Glu Leu Gly Phe Ser Ser Asn Leu Leu
177         35                   40                  45
179 tgc agc tct tgt gat ctt ctc gga cag ttc aac ctg ctt cag ctg gat 192
180 Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe Asn Leu Leu Gln Leu Asp
181     50                  55                  60
183 cct gat tgc aga gga tgc tgt cag gag gaa gca caa ttt gaa acc aaa 240
184 Pro Asp Cys Arg Gly Cys Cys Gln Glu Glu Ala Gln Phe Glu Thr Lys
185  65                  70                  75                  80
187 aag ctg tat gca gga gct att ctt gaa gtt tgt gga tga aaa ttg gga 288
188 Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val Cys Gly Xaa Lys Leu Gly
189                 85                  90                  95
191 agg ttc cct caa gtc caa gct ttt gtt agg agt gat aaa ccc aaa ctg 336
192 Arg Phe Pro Gln Val Gln Ala Phe Val Arg Ser Asp Lys Pro Lys Leu
193             100                 105                 110
195 ttc aga gga ctg caa atc aag tat gtc cgt ggt tca gac cct gta tta 384
196 Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg Gly Ser Asp Pro Val Leu
197         115                 120                 125
199 aag ctt ttg gac gac aat ggg aac att gct gaa gaa ctg agc att ctc 432
200 Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala Glu Glu Leu Ser Ile Leu
201     130                 135                 140
203 aaa tgg aac aca gac agt gta gaa gaa ttc ctg agt gaa aag ttg gaa 480
204 Lys Trp Asn Thr Asp Ser Val Glu Glu Phe Leu Ser Glu Lys Leu Glu
205 145                 150                 155                 160
207 cgc ata taa                                                      489
208 Arg Ile
211 <210> SEQ ID NO: 4
212 <211> LENGTH: 136
```

W--> 188

RAW SEQUENCE LISTING   DATE: 10/19/2000
PATENT APPLICATION: US/09/676,718   TIME: 11:24:31

Input Set : A:\56113.app
Output Set: N:\CRF3\10192000\I676718.raw

```
213 <212> TYPE: PRT
214 <213> ORGANISM: Homo sapiens
216 <220> FEATURE:
217 <221> NAME/KEY: SITE
218 <222> LOCATION: (67)
219 <223> OTHER INFORMATION: Xaa is selenocysteine
221 <400> SEQUENCE: 4
222 Ser Ala Phe Gly Ala Glu Phe Ser Ser Glu Ala Cys Arg Glu Leu Gly
223   1               5                   10                  15
225 Phe Ser Ser Asn Leu Leu Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe
226                 20                  25                  30
228 Asn Leu Leu Gln Leu Asp Pro Asp Cys Arg Gly Cys Cys Gln Glu Glu
229             35                  40                  45
231 Ala Gln Phe Glu Thr Lys Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val
232         50                  55                  60
234 Cys Gly Xaa Lys Leu Gly Arg Phe Pro Gln Val Gln Ala Phe Val Arg
235   65                  70                  75                  80
237 Ser Asp Lys Pro Lys Leu Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg
238                 85                  90                  95
240 Gly Ser Asp Pro Val Leu Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala
241                 100                 105                 110
243 Glu Glu Leu Ser Ile Leu Lys Trp Asn Thr Asp Ser Val Glu Glu Phe
244             115                 120                 125
246 Leu Ser Glu Lys Leu Glu Arg Ile
247         130                 135
250 <210> SEQ ID NO: 5
251 <211> LENGTH: 21
252 <212> TYPE: DNA
253 <213> ORGANISM: Artificial Sequence
255 <220> FEATURE:
256 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR primer
258 <400> SEQUENCE: 5
259 atggcggctg ggccgagtgg g                                  21
262 <210> SEQ ID NO: 6
263 <211> LENGTH: 21
264 <212> TYPE: DNA
265 <213> ORGANISM: Artificial Sequence
267 <220> FEATURE:
268 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR primer
270 <400> SEQUENCE: 6
271 taatatgcgt tccaactttt c                                  21
274 <210> SEQ ID NO: 7
275 <211> LENGTH: 21
276 <212> TYPE: DNA
277 <213> ORGANISM: Artificial Sequence
279 <220> FEATURE:
280 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR primer
282 <400> SEQUENCE: 7
283 tctgcttttg gggcagagtt t                                  21
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/676,718

DATE: 10/19/2000
TIME: 11:24:31

Input Set : A:\56113.app
Output Set: N:\CRF3\10192000\I676718.raw

```
286 <210> SEQ ID NO: 8
287 <211> LENGTH: 1216
288 <212> TYPE: DNA
289 <213> ORGANISM: Mus musculus
291 <220> FEATURE:
292 <221> NAME/KEY: CDS
293 <222> LOCATION: (11)..(490)
295 <220> FEATURE:
296 <221> NAME/KEY: misc_feature
297 <222> LOCATION: (287)..(289)
298 <223> OTHER INFORMATION: TGA codon codes for Selenocysteine, Xaa
300 <400> SEQUENCE: 8
301 gaccgcaggg atg gcg gca ggg cag ggt ggg tgg ctg cgg cca gct ctg    49
302             Met Ala Ala Gly Gln Gly Gly Trp Leu Arg Pro Ala Leu
303              1               5                  10
305 ggg ctg cgc ttg ctg ctg gcg act gcg ttt caa gcg gtg tct gct ctg    97
306 Gly Leu Arg Leu Leu Leu Ala Thr Ala Phe Gln Ala Val Ser Ala Leu
307      15                  20                  25
309 ggg gca gag ttt gcg tca gag gca tgc aga gag ttg ggt ttc tcc agc   145
310 Gly Ala Glu Phe Ala Ser Glu Ala Cys Arg Glu Leu Gly Phe Ser Ser
311  30                  35                  40                  45
313 aac ttg ctc tgc agc tct tgc gat ctt ctt gga cag ttt aat ctg ctc   193
314 Asn Leu Leu Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe Asn Leu Leu
315                  50                  55                  60
317 cca ctg gac cct gtt tgc aga ggg tgc tgt cag gaa gaa gca caa ttt   241
318 Pro Leu Asp Pro Val Cys Arg Gly Cys Cys Gln Glu Glu Ala Gln Phe
319                  65                  70                  75
321 gaa acc aaa aag ctg tat gca gga gcc atc ctt gaa gtc tgc gga tga   289
322 Glu Thr Lys Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val Cys Gly Xaa
323              80                  85                  90
325 aaa ttg ggg agg ttc cct caa gtc caa gct ttt gtc aga agt gat aaa   337
326 Lys Leu Gly Arg Phe Pro Gln Val Gln Ala Phe Val Arg Ser Asp Lys
327      95                  100                 105
329 ccc aaa ctc ttc aga ggt cta cag atc aag tat gtt cga ggc tca gac   385
330 Pro Lys Leu Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg Gly Ser Asp
331 110                 115                 120                 125
333 cct gta cta aag ctt ttg gac gac aac ggg aac att gct gaa gaa cta   433
334 Pro Val Leu Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala Glu Glu Leu
335                 130                 135                 140
337 agc atc ctc aaa tgg aac aca gac agt gtg gaa gag ttc ctg agc gag   481
338 Ser Ile Leu Lys Trp Asn Thr Asp Ser Val Glu Glu Phe Leu Ser Glu
339             145                 150                 155
341 aag ttg gaa cgcatataaa catgcttagt agttttata ctaatcaaat           530
342 Lys Leu Glu
343             160
345 gaattatcac agcacctaga caataactta gttttgcatg cttacattgg tcatcctttt 590
347 tatgtacatc attaatcttc tgacaagaag ctgaagtagc accacagtcc ataatatatc 650
349 aggatctggc aagcttaagg aacccagctc ttagaaattt ctcttcttct acacttgttg 710
351 ctctcaccag tgaaacgctt tgtaaggagg catctgggta attatgcaaa taagtttgtg 770
```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY   DATE: 10/19/2000
PATENT APPLICATION: US/09/676,718   TIME: 11:24:32

Input Set : A:\56113.app
Output Set: N:\CRF3\10192000\I676718.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:49 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:104 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:188 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:234 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:322 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:396 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9